

CAIXA: a Catalogue of AGN In the XMM-Newton Archive

I. Spectral analysis

Stefano Bianchi^{1,2}, Matteo Guainazzi², Giorgio Matt¹, Nuria Fonseca Bonilla², Gabriele Ponti³

¹ Dipartimento di Fisica, Università degli Studi Roma Tre, via della Vasca Navale 84, 00146 Roma, Italy

² XMM-Newton Science Operations Center, European Space Astronomy Center, ESA, Apartado 50727, E-28080 Madrid, Spain

³ Laboratoire APC, UMR 7164, 10 rue A. Domon et L. Duquet, 75205 Paris, France

Received / Accepted

ABSTRACT

Aims. We present CAIXA, a Catalogue of AGN In the XMM-Newton Archive. It consists of all the radio-quiet X-ray unobscured ($N_H < 2 \times 10^{22} \text{ cm}^{-2}$) Active Galactic Nuclei (AGN) observed by XMM-Newton in targeted observations, whose data are public as of March 2007. With its 156 sources, this is the largest catalogue of high signal-to-noise X-ray spectra of AGN.

Methods. All the EPIC pn spectra of the sources in CAIXA were extracted homogeneously and a baseline model was applied in order to derive their basic X-ray properties. These data are complemented by multiwavelength data found in the literature: Black Hole masses, Full Width Half Maximum (FWHM) of H β , radio and optical fluxes.

Results. Here we describe our homogeneous spectral analysis of the X-ray data in CAIXA and present all the results on the parameters adopted in our best-fit models.

Key words. Galaxies: active - Galaxies: Seyfert - quasars: general - X-rays: general

1. Introduction

Despite the unquestionable progress made in our understanding of the physics of Active Galactic Nuclei (AGN), there are still many open issues related to the X-ray properties of these objects. Some of the most discussed ones in the last years are: the highly debated nature of the soft X-ray excess in unobscured sources; the different spectral and timing properties for different classes of sources; the nature of radio emission in radio-quiet objects and its relation to the X-ray nuclear emission; the origin of the reprocessing of the primary emission from highly ionized material; the correlation of the above-mentioned phenomenology with fundamental properties of AGN, such as the Black Hole (BH) mass and the accretion rate. The most effective way to address these questions is to analyze large numbers of AGN with good-quality X-ray spectra and to perform statistical studies, taking into account data in other wavelengths and other basic properties of the objects.

The European Space Agency's (ESA) X-ray Multi-Mirror Mission (XMM-Newton) was launched on December 10th 1999. The XMM-Newton public archive has since become the repository of an enormous amount of high-quality X-ray data and a precious legacy for future missions. In particular, our knowledge of the physics of AGN has dramatically improved thanks to the large sensitivity of the European Photon Imaging Camera (EPIC) pn charge-coupled device (CCD) arrays. At this time, a systematic and homogeneous study of the EPIC pn spectra of AGN represents a necessary step to fully take advantage of this highly successful X-ray mission.

In this paper, we present CAIXA, a Catalogue of AGN In the XMM-Newton Archive. This catalogue was already used by Bianchi et al. (2007) to confirm with much higher confidence the

'Iwasawa-Taniguchi' effect (i.e. the anti-correlation between the Equivalent Width of the neutral iron narrow emission line and the X-ray luminosity) and by Guainazzi et al. (2006a, a smaller catalogue was used at that period) and Longinotti et al. (2007), to assess the frequency of the relativistic component of the iron line. Here we describe our homogeneous spectral analysis of the X-ray data in CAIXA and present all the results on the parameters adopted in our best-fit models. In following papers, we will present the timing analysis and properties of CAIXA and we will investigate the correlations between the X-ray and the multiwavelength properties of the sources in the catalogue.

2. The catalogue

2.1. X-ray data reduction and spectral analysis

CAIXA consists of all the radio-quiet X-ray unobscured ($N_H < 2 \times 10^{22} \text{ cm}^{-2}$) AGN observed by XMM-Newton in targeted observations, whose data are public as of March 2007. Only EPIC pn (Strüder et al. 2001) data were reprocessed, with SASv6.5 (Gabriel et al. 2004). For the observations performed in Small Window mode, background spectra were generated using blank-field event lists, according to the procedure presented in Read & Ponman (2003). In all other cases, background spectra were extracted from source-free regions close to the target in the observation event file.

Source extraction radii and screening for intervals of flaring particle background were performed via an iterative process which leads to a maximization of the signal-to-noise ratio, similarly to what described in Piconcelli et al. (2004). Spectra were binned in order to oversample the intrinsic instrumental energy resolution by a factor not lower than 3, and to have spectral bins with at least 25 background-subtracted counts. This ensures the applicability of the χ^2 statistics.

We applied a number of criteria to filter the catalogue. First of all, spectra with less than 200 counts in either of the (rest-frame) bands of 0.5-2 and 2-10 keV were rejected, because they do not possess enough independent bins to be fitted with our models. Moreover, spectra affected by a pileup larger than 1% were rejected. Only two objects, namely NGC 2992 and ESO 548-G081, did not have an alternative pileup-free observation and were therefore excluded from the catalogue. After the fitting procedure, all sources with a local column density (measured in the 2-10 keV band) larger than $2 \times 10^{22} \text{ cm}^{-2}$ were further excluded from the catalogue.

In order to ensure that only radio-quiet object were included in CAIXA, we collected radio data at 6 cm (4.85 GHz) and 20 cm (1.4 GHz) from the literature and calculated the K -corrected radio-loudness parameter R ($\log R = \log f_{6\text{cm}} - \log f_B$; Stocke et al. 1992) and the X-ray radio-loudness parameter R_X ($\log R_X = \log f_{6\text{cm}} - \log f_{2-10\text{keV}}$; Terashima & Wilson 2003). When only a 20 cm radio flux was available, it was converted to 6 cm, assuming a power law radio spectrum ($f_\nu \propto \nu^{-\alpha_r}$), where $\alpha_r = 0.7$ is the average two-point radio spectral index for the sources in CAIXA with measures both at 6 and 20 cm (61 sources). The same procedure was followed for the (few) 843 MHz measures. The optical flux f_B was derived from the V magnitude and the colour index $B-V$, the latter assumed to be 0.3 if not available in literature. According to the standard definition, all quasars with $\log(R) > 1$ were excluded, while for Seyfert galaxies, known to be on average ‘radio-louder’, $\log(R) > 2.4$ and $\log(R_X) > -2.755$ were used as boundaries (see Panessa et al. 2007). Note that three quasars are nominally radio-loud, having values of R of 1.09 (RXJ0136.9-3510), 1.01 (RXJ0228.2-4057) and 1.08 (KUV18217+6419). However, given that the uncertainties on the radio and optical fluxes are not taken into consideration, we are confident these sources can be safely considered radio-quiet. Moreover, we did not find radio measures for the following sources: ESO 198-G24, RXJ0323.2-4931, Fairall 1116, 1H0707-495, J124126.3-574955, H1846-786, RXJ2241.8-4405, REJ2248-511. Some of them are treated as radio-quiet sources in papers cited in Table 3 and, in any case, the vast majority of them are Seyfert galaxies and unlikely to be radio-loud with our criteria. Therefore, we kept all of them in our final catalogue. Finally, with respect to the catalogue used in Bianchi et al. (2007), we excluded PMNJ0623-6436, for which we found a new radio observation at 6 cm (Sadler et al. 2006), leading to $\log R \simeq 2$. Nevertheless, this exclusion does not affect the results published in Bianchi et al. (2007).

At the end of this selection procedure, the total catalogue comprises 156 radio-quiet AGN. When more than one observation of the same object was available, we used the one with the longest exposure, independently of its flux status, in order to avoid, as much as possible, any bias. Tables 1 and 2 summarize the properties of the XMM-Newton observations included in CAIXA, along with the fluxes, luminosities and main X-ray parameters derived from the fits with the baseline model.

The baseline model consists of the following components:

$$F(E) = e^{-\sigma(E)N_H^G} \left[A e^{-\sigma(E)N_H^s} E^{-\Gamma_s} + B e^{-\sigma(E)N_H^h} E^{-\Gamma_h} + 4 \times G \right] \quad (1)$$

where $\sigma(E)$ is the photoelectric cross-section, adopting solar abundances as in Anders & Ebihara (1982), N_H^G is the Galactic column density appropriate for each source (after Dickey & Lockman 1990), $N_H^s < N_H^h$ are two local column densities at the redshift of the source, possibly coexisting if related to absorption on different scales, A and B are two normalization factors, Γ_s and Γ_h the spectral photon indexes for the soft

an the hard spectrum, and G are four Gaussian emission lines at rest-frame energies fixed to 6.4, 6.7, 6.96 and 7.06 keV, as appropriate for neutral, He- and H-like iron $K\alpha$ and neutral iron $K\beta$, respectively. The inclusion of the latter allows us to disentangle its possible contamination to the 6.96 keV line, by not allowing its flux to be larger than 0.16 times the flux of the neutral $K\alpha$ line (Molendi et al. 2003). A relativistically broadened component of the iron line was never included in our fits (see Bianchi et al. 2007, for a discussion on this point). Where the addition of the second powerlaw leads to a χ^2 improvement less than 99 % (according to F-test), we put $B = 0$ and the baseline model only includes a powerlaw component. Finally, two absorption edges at 0.737 (O VII) and 0.871 (O VIII) keV are included when required, to roughly model a warm absorber (see also Sect. 3.1 and 3.2 for a discussion on this issue). No Compton reflection component is included in our fits, since the limited bandpass of XMM-Newton does not allow us to give good constraints on this parameter, in particular when the SNR is not very large. We will discuss the possible consequences of this choice on our results in Sect. 3.2. In order to study the correlations where the soft X-ray luminosities are involved, we decided to compile a sub-catalogue, where all the sources with a column density in excess of the Galactic one (but still lower than the general $2 \times 10^{22} \text{ cm}^{-2}$ limit of CAIXA) were excluded. This sub-catalogue comprises 142 sources and will be explicitly referred to when used in this paper instead of the whole CAIXA.

All spectra were analyzed with XSPEC 12.2.1 (Arnaud 1996). In the following, errors correspond to the 90% confidence level for one interesting parameter ($\Delta\chi^2 = 2.71$), where not otherwise stated. The cosmological parameters used throughout this paper are $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.73$ and $\Omega_m = 0.23$ (i.e. the default ones in XSPEC 12.2.1).

2.2. Multiwavelength data and BH masses

Table 3 summarizes the multiwavelength data collected for the sources in CAIXA, i.e. the absolute magnitude (100% coverage), the radio fluxes at 6 and/or 20 cm (95% coverage), Full Width Half Maximum (FWHM) of $H\beta$ (65% coverage), BH mass (57%), V magnitude (100% coverage) and $B-V$ colour index (65% coverage). All the objects of the catalogue are included in Veron-Cetty & Veron (2006), who represented the main reference for the multiwavelength data. All the other references are shown in Table 3. We stress here that the FWHM of $H\beta$ in CAIXA are all larger than 650 km s^{-1} , as expected since it includes only sources unobscured or moderately obscured in the X-rays. Therefore, our distinction between narrow- and broad-line objects simply reflects the standard limit of $H\beta \text{ FWHM} < 2000 \text{ km s}^{-1}$, while no optical Type 2 objects are present in CAIXA. Moreover, the standard boundary of $M_{\text{abs}} = -23$ was adopted to classify a source as a quasar or as a Seyfert galaxy.

2.3. General properties

CAIXA includes a total of 77 quasars (16 narrow-line and 35 broad-line objects) and 79 Seyfert galaxies (21 narrow-line and 30 broad-line objects). The redshift distribution, which spans from $z=0.002$ to $z=4.520$, is shown in Fig. 1: given the requirement of a high SNR X-ray spectrum, most of the sources are local (almost 90% within $z < 1$). Also shown in the same figure are the BH masses distribution ($1.4 \times 10^6 M_\odot < M_{\text{BH}} < 2.0 \times 10^{10} M_\odot$): as expected, quasars have, on average, larger BH

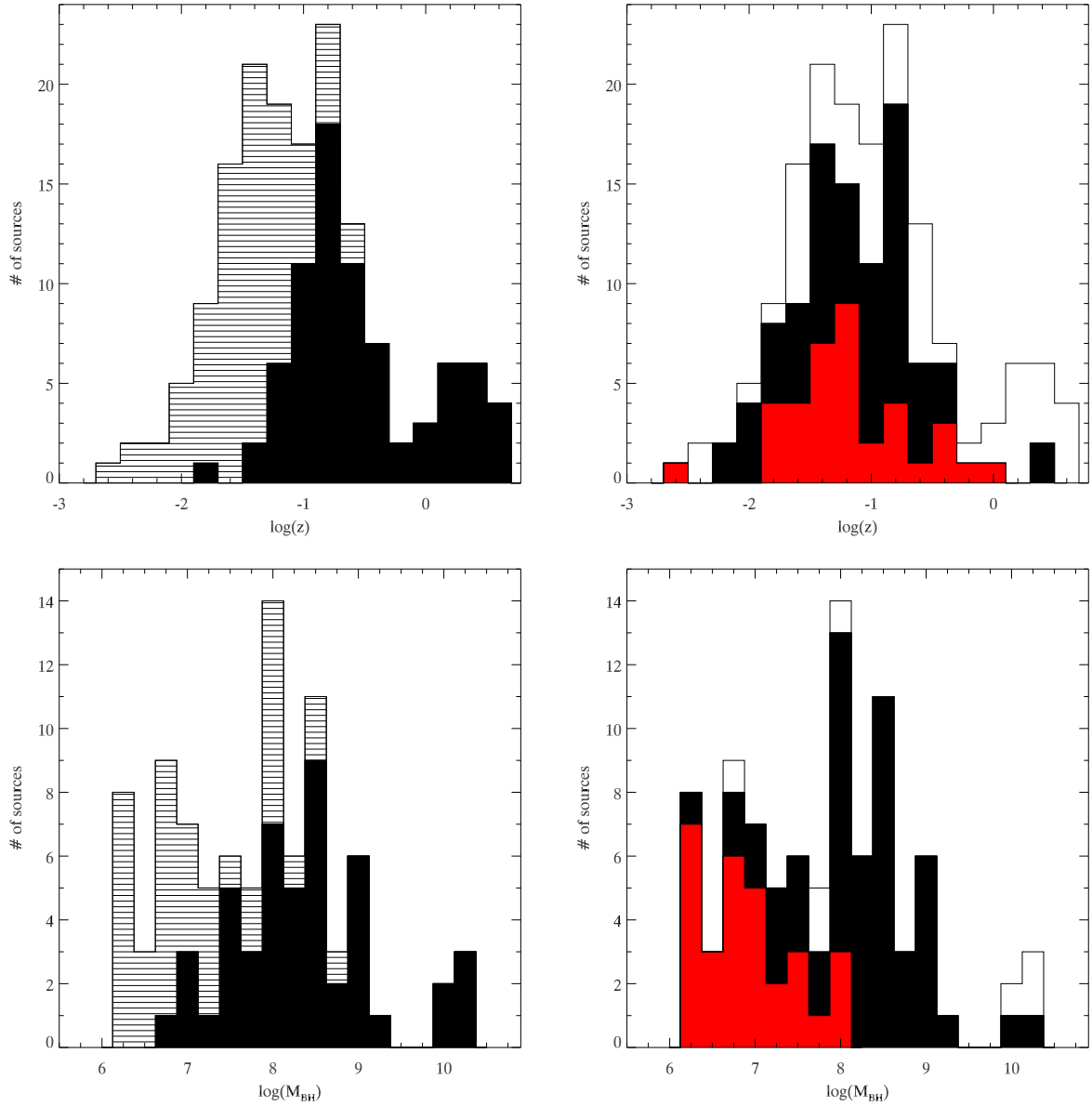


Fig. 1. General properties of the sources included in CAIXA. Redshift (*top*) and BH mass (*bottom*) distributions. Quasars (filled) and Seyferts (dashed) are shown on the left, while narrow-line objects (red), broad-line objects (black) and objects without a measure of the $H\beta$ FWHM (white) are shown on the right. See text for details.

masses and narrow-line objects the smallest ones, given the correlation between the BH mass and the FWHM of the $H\beta$.

The distribution of the hard X-ray luminosities, which cover a range between $L_{2-10\text{keV}} = 2.0 \times 10^{41}$ to $3.9 \times 10^{46} \text{ erg s}^{-1}$, is shown in Fig. 2. The customary X-ray boundary between quasars and Seyferts at $L_{2-10\text{keV}} = 1 \times 10^{44} \text{ erg s}^{-1}$ seems to be in fair agreement with the historical optical classification adopted in CAIXA, even if there is a larger number of X-ray weak quasars. Finally, the distribution of the Eddington ratios are shown in the lower panels of Fig. 2: these values are derived from the 2-10 keV luminosities, adopting a luminosity-dependent bolometric correction, after Marconi et al. (2004). The average Eddington ratio is $\log(L_{\text{bol}}/L_{\text{Edd}}) = -0.79 \pm 0.08$ (standard deviation $\sigma = 0.71$) for all the catalogue, -0.69 ± 0.08 ($\sigma = 0.55$) for quasars, -0.90 ± 0.13 ($\sigma = 0.86$) for Seyferts, -0.56 ± 0.13 ($\sigma = 0.73$) for narrow-line objects and -0.84 ± 0.08 ($\sigma = 0.58$) for broad-line

objects. These values are in agreement with other samples and catalogues (e.g. McLure & Dunlop 2004; Bian & Zhao 2004; Netzer & Trakhtenbrot 2007), even if in CAIXA narrow-line objects do not show, on average, very large values.

3. Results

3.1. Goodness of fits

Fig. 3 shows the χ^2 values obtained from the automatic fit procedure of CAIXA as a function of the degrees of freedom (dof) of each observation. The presence of a second powerlaw to model the soft X-ray emission is statistically required in 123 spectra out of 156 (see also Sect. 2.1): the single-powerlaw fits are marked in orange in Fig. 3. As expected, they are clustered in the low-dof part of the plot. The region where the adopted model can be

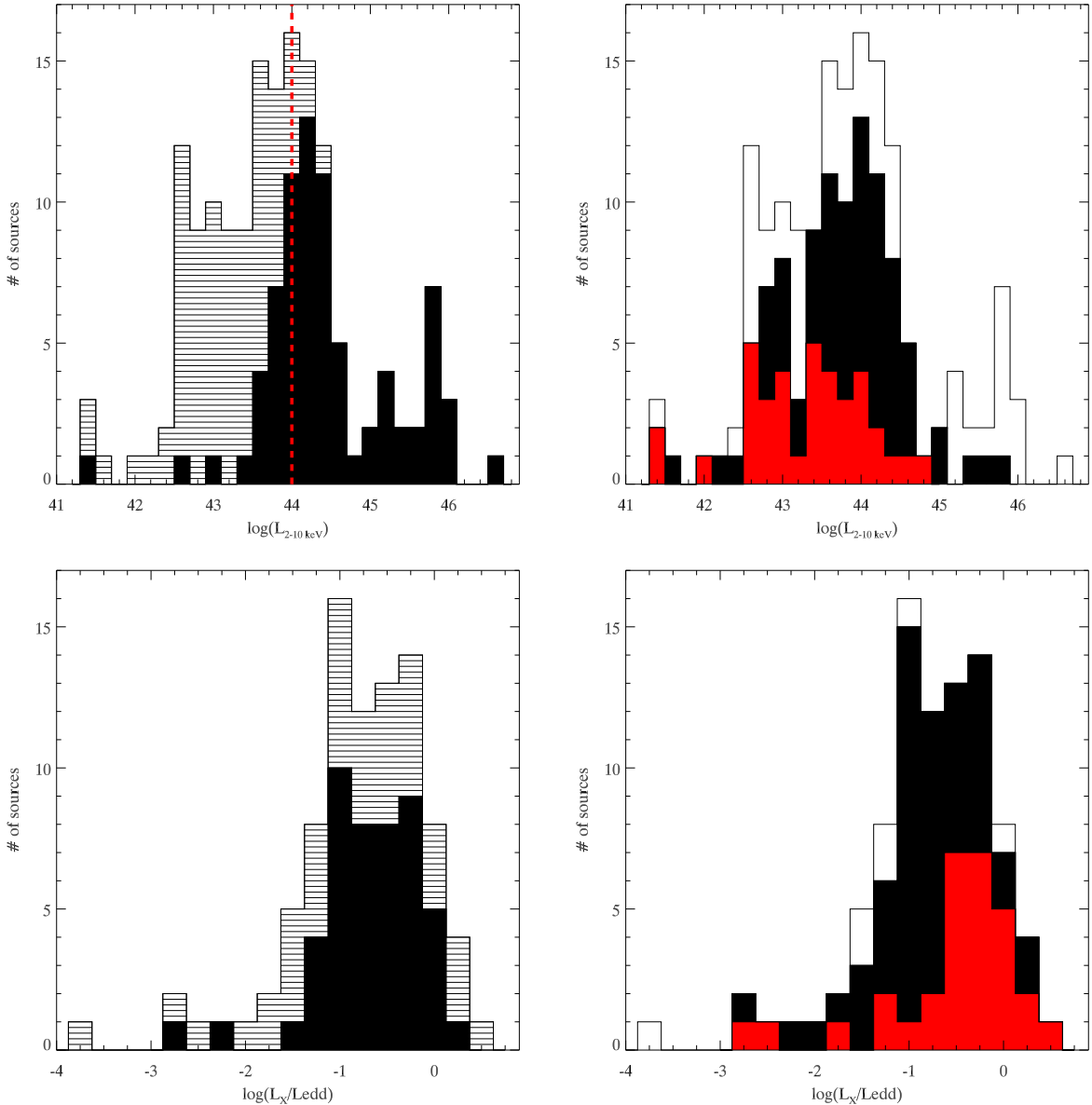


Fig. 2. General properties of the sources included in CAIXA. Hard X-ray (2-10 keV) luminosity (*top*) and Eddington ratio (*bottom*) distributions. Quasars (filled) and Seyferts (dashed) are shown on the left, while narrow-line objects (red), broad-line objects (black) and objects without a measure of the H β FWHM (white) are shown on the right. See text for details.

considered correct at the 99% confidence level is also shown. In other words, for each dof value there is a 1% probability of getting a value of χ^2 as large or larger than the dot-dashed curve plotted in figure if the model is correct. Most of the fits in CAIXA lie under this curve, confirming the good quality of the derived spectral parameter. As an example, we show in Fig. 4 the best fit for Mrk 590 ($\chi^2=228/229$), which, with its 229 dof (the maximum value allowed by the spectral binning criteria described in Sect. 2.1) is one of the observations with the highest SNR.

However, there is a number of fits whose χ^2 cannot be considered statistically acceptable. Not surprisingly, most of them correspond to very large dof values: high SNR spectra are more difficult to be properly fitted by the simple model we are using in CAIXA. As an example, we show in Fig. 5 the “best fit” for NGC 3783 ($\chi^2=2268/229$), which, together with MCG-6-30-15,

is one with the poorest reduced χ^2 in CAIXA. As apparent from the left panel of Fig. 5, the main problem is represented by huge residuals in the soft X-ray part of the spectrum, where the effects of absorption from ionised matter is only roughly taken into account by two absorption edges (see Sect. 2.1).

The lack of a proper modeling of warm absorbers in CAIXA is dictated by the complexity of these components, which do not allow an automatic handling. But, since we are not extracting information on this component in our catalogue, we should only check how much the spectral properties we are interested in are affected by this issue. In order to verify that, we fitted the NGC 3783 data with the model adopted by Reeves et al. (2004), which includes three phases for the warm absorbers, a thermal black body for the soft excess, a powerlaw and a Compton reflection component for the hard X-rays and the iron lines. The resulting χ^2 is acceptable (303/221) and compara-

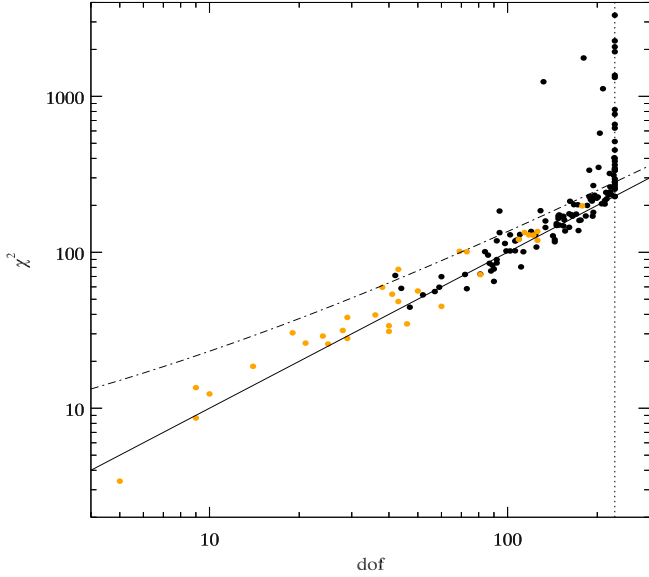


Fig. 3. χ^2 values obtained from the automatic fit procedure of CAIXA as a function of the degrees of freedom (dof) of each observation. Orange points are those where the second power-law is not statistically required. The continuous line refers to $\chi^2/\text{dof}=1$, while the dot-dashed curve the χ^2 value corresponding to a 99% Null Hypothesis Probability for the corresponding dof value. The vertical dotted line represents the 229 dof limit derived from the spectral binning criteria in CAIXA. See text for details.

ble with the one found by Reeves et al. (2004). The 0.5-2 keV flux is $16.06 \pm 0.02 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$, while the 2-10 keV flux is $55.29 \pm 0.06 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$. These values should be compared to the ones automatically calculated in CAIXA: $15.34 \pm 0.12 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$, while the 2-10 keV flux is $55.48 \pm 0.15 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$. Therefore, the difference in the soft X-ray flux is lower than 5%, while the hard X-ray fluxes are consistent, within errors, and their nominal difference is of the order of 3%. The EWs of the three iron lines are all consistent with the results obtained in CAIXA: 91^{+5}_{-4} (93^{+4}_{-5}) eV, < 1 (< 1) eV and 25 ± 5 (17^{+3}_{-6}) eV, respectively for neutral iron, Fe XXV and Fe XXVI.

We therefore conclude that, even when the statistical significance of the best fit in CAIXA is very low, as in the case presented above, the values of the X-ray parameters derived from the fit are still very reliable and can be safely used to properly characterise the sources. The only parameter which can be significantly affected by the simplicity of the model adopted in CAIXA is the photon spectral index. In the case of NGC 3783, for example, the value derived from the automatic fit procedure is very flat, being 1.49. The addition of a Compton reflection component (with $R=1$, consistent with the neutral iron $K\alpha$ line, as suggested by Reeves et al. (2004)) and a complex warm absorber significantly steepens the powerlaw index to a more canonical value, 1.76. We will discuss this issue in the next section.

3.2. The X-ray spectral index

The average¹ 2-10 keV spectral index for the whole catalogue is 1.73 ± 0.04 , with a large spread of $\sigma = 0.45 \pm 0.03$. The hard X-ray Γ appears steeper in quasars (1.80 ± 0.05 , consistent with the value obtained for the PG quasars: Piconcelli et al. 2005) than in Seyferts (1.66 ± 0.05). A Kolmogorov-Smirnov test yields that the two populations are different at the 99% confidence level. The difference is even larger if we consider narrow-line objects ($\Gamma = 1.94 \pm 0.07$) with respect to broad-line objects (1.62 ± 0.04). In this case, a K-S test gives a NHP of 2×10^{-5} : the two populations are significantly different with respect to their X-ray photon indexes. It is important to stress here that the photon indexes quoted in CAIXA, in principle, are not the intrinsic ones, because in our fits we do not consider the effect of Compton reflection and warm absorption. Therefore, the reader should keep in mind that the intrinsic photon indexes are likely to be steeper than the ones reported here, as in the case presented in the previous section. Indeed, when the Compton reflection component can be well constrained, as in broad band *BeppoSAX* spectra, the average spectral index is steeper (1.89 ± 0.03 for Seyfert galaxies: Dadina 2008). It is, therefore, possible that the different photon indexes found for different populations in CAIXA do not only (or at all) reflect differences in the intrinsic Γ of the sources, but also on the amount of Compton reflection.

In order to test this hypothesis, we performed a second run of automating fitting procedure, with a model which includes Compton reflection. Since the amount of reflection cannot be constrained with the EPIC pn band coverage (and this is the reason why it is not included in the default model adopted in CAIXA) we fixed it in each source to a value tightly linked with the EW of the neutral iron $K\alpha$ line. A linear correlation between the two parameters was assumed, adopting $R = \frac{\Omega}{2\pi} = 1$ for $\text{EW}=120$ eV as normalization (see e.g. Matt et al. 1991). Furthermore, in order to avoid the technical issues involved in a fit with R linked to the EW of an emission line, we derived the expected iron $K\alpha$ EW on the basis of the hard X-ray luminosity, following the best fit for the Iwasawa-Taniguchi effect (Bianchi et al. 2007):

$$\log(\text{EW}_{\text{Fe}}) = (1.73 \pm 0.03) + (-0.17 \pm 0.03) \log(L_{\text{X},44}) \quad (2)$$

where EW_{Fe} is the EW of the neutral iron $K\alpha$ line in eV and $L_{\text{X},44}$ is the 2-10 keV X-ray luminosity in units of $10^{44} \text{ erg s}^{-1}$. We, therefore, introduce a luminosity-dependent Compton reflection component. As for the default model in CAIXA, we added a second powerlaw to model the soft X-ray emission, only if statistically required.

The distributions of the hard X-ray powerlaw index derived for the model which includes Compton reflection are shown in the bottom panels of Fig. 6. The average 2-10 keV spectral index for the whole catalogue is, as expected, steeper, but not significantly, being now 1.78 ± 0.04 ($\sigma = 0.46 \pm 0.03$). Quasars (1.83 ± 0.05) and Seyferts (1.73 ± 0.05) are still different at the 99% level, on the basis of a K-S test. Narrow-line ($\Gamma = 2.00 \pm 0.07$) and broad-line objects (1.65 ± 0.04) still display a very significant difference (NHP of 4×10^{-5} for the K-S test). The inclusion of a luminosity-dependent Compton reflection component, therefore, does not significantly affect the distribution of the powerlaw indexes in CAIXA and, more importantly, does not alter the differences between populations found with the simpler model.

¹ These mean values are ‘censored’, because they take into account errors and upper limits on the measures, using the method described in Appendix A.

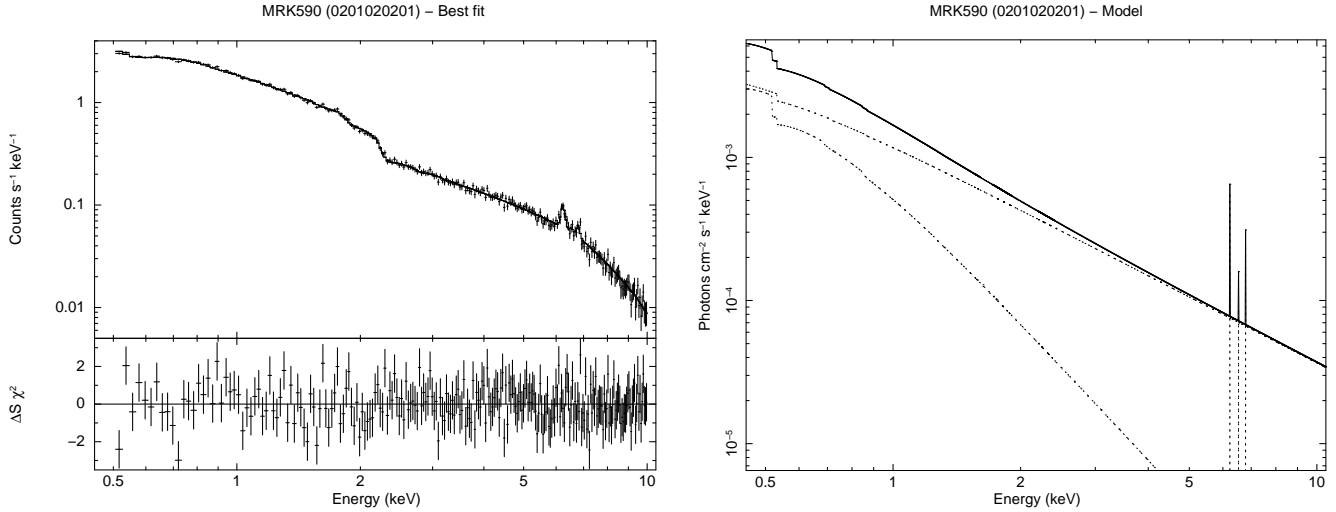


Fig. 4. Best fit data and model for a good fit in CAIXA, Mrk 590. Spectral parameters derived from this automatic fit are statistically reliable.

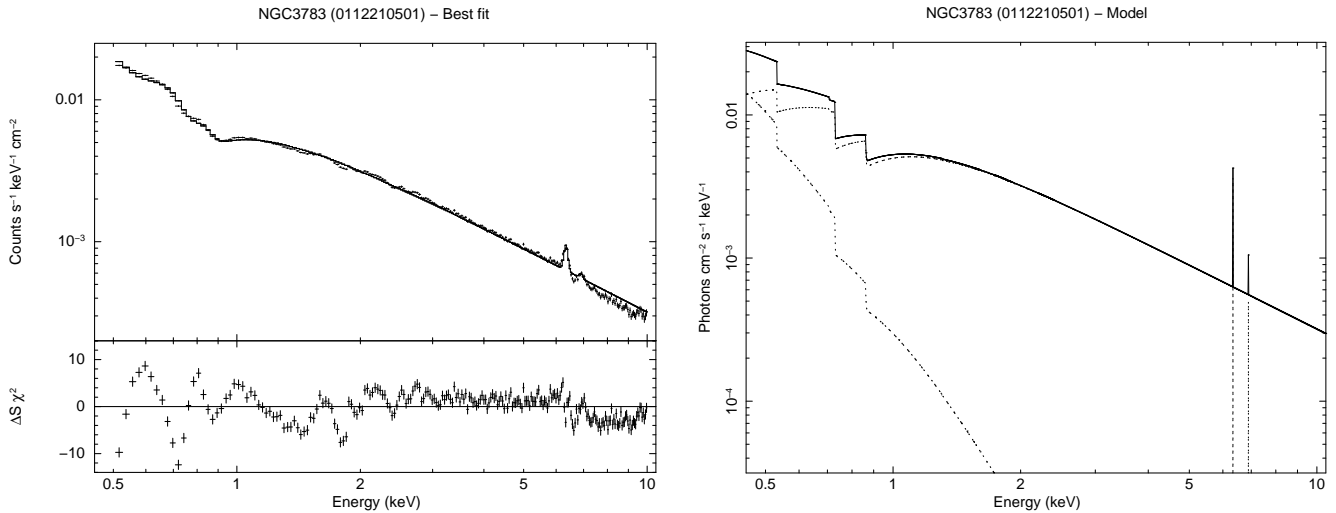


Fig. 5. Best fit data and model for a bad fit in CAIXA, NGC 3783. The spectral parameters derived from this automatic fit are still consistent with those extracted from a proper modeling of the data. See text for details.

Another important component which is not included in the default model adopted in CAIXA is the warm absorber, which is known to be present in many sources (see e.g. Piconcelli et al. 2004). If it is not properly modeled, it may result in a flatter powerlaw index, as in the case of NGC 3783 reported in Sect. 3.1. However, a correct modelisation of the complex (and often multiple) warm absorption component is far beyond the scopes of CAIXA and will not be addressed any more in the following. Nevertheless, we would like to stress that most of the studies on large samples of AGN adopt simple models, so that the hard X-ray powerlaw index reported in CAIXA can be directly compared to the ones found in literature. On the other hand, it is clear that the value of the intrinsic Γ for single sources may significantly differ from those found with more complex models, but this is inherent in any statistical analysis of catalogues and the effect is diluted by the large numbers of sources.

3.3. The iron lines

The model adopted to fit the X-ray spectra in CAIXA includes the $K\alpha$ emission lines from neutral iron, Fe XXV and Fe XXVI (see Sect. 2.1). In Fig. 7, we show the distributions of the EW for the three iron lines. The neutral iron line is detected in 81 sources, 46 of which are broad-line objects (out of the 65 in CAIXA) and 18 narrow-line objects (out of the 37 in CAIXA)². The distribution is clustered around 50-120 eV, as expected if arising as reprocessing from cold Compton-thick matter, as the putative torus (see e.g. Matt et al. 1991). Only 4 sources have (detected) EWs larger than 250 eV, while the average (including upper limits) value of the EW is 76 ± 6 eV, with quite a large spread ($\sigma = 75$ eV). Such a large spread is also expected, since the neutral iron $K\alpha$ EW is strongly anti-correlated to the X-ray luminosity (the Iwasawa-Taniguchi effect: Bianchi et al. 2007).

² These values are slightly different from the ones reported in Bianchi et al. (2007), being 81, 43 and 19, because of the update of the $H\beta$ FWHM values in CAIXA.

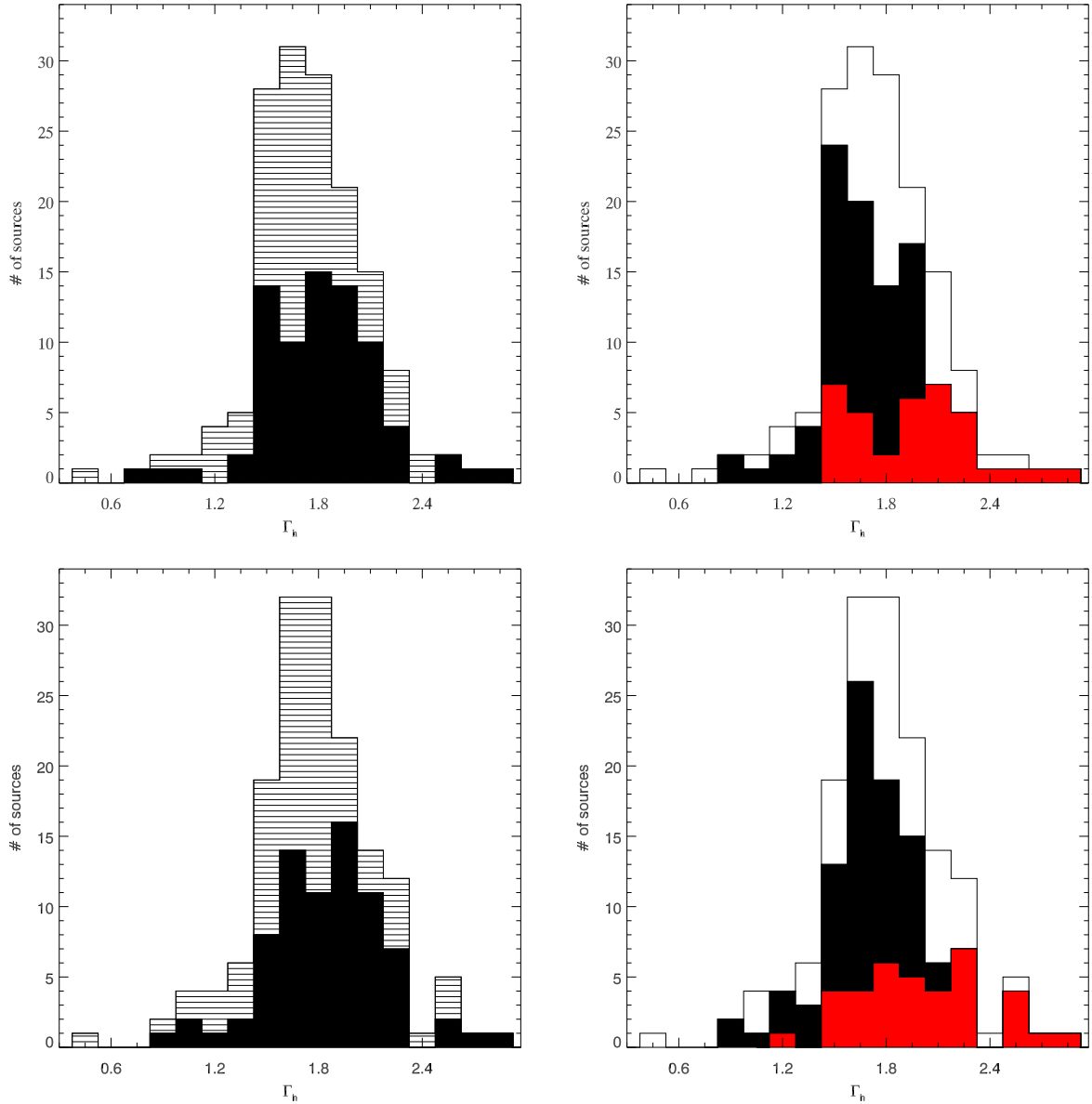


Fig. 6. Hard X-ray powerlaw index distribution in CAIXA, for the best fit model (*upper panels*) and the model which includes a luminosity-dependent Compton reflection component (*lower panels*). On the left, dashed histograms are Seyfert galaxies, filled ones are quasars. On the right, red histograms are narrow-line objects, black histograms are broad-line objects, while white histograms are used for sources without a measure of the H β FWHM in CAIXA. See text for details.

The Fe XXV line is detected in 44 sources, 26 of which are broad-line objects and 11 narrow-line objects (out of the 37 in CAIXA). Similarly, the Fe XXVI line is detected in 40 sources, 21 broad-line objects and 12 narrow-line objects. We therefore do not find different rates of detection of ionised iron lines for these two classes of AGN. These lines may be produced in a Compton-thin gas photoionised by the AGN, but EWs larger than ≈ 100 eV are quite difficult to explain within this scenario (see e.g. Bianchi & Matt 2002; Bianchi et al. 2005). The effect of the neutral K β line which may, in principle, contaminate the flux of the Fe XXVI line, is properly taken into account in CAIXA. However, note that lines with such a large value of EW are few in CAIXA and always with very large errors (see Fig. 7, considering that EWs larger than 130 eV are detected only in 4 and 2 sources, for Fe XXV and Fe XXVI, respectively).

In any case, we cannot exclude that an ionised accretion disc may contribute to the flux of the ionised iron lines in the objects where the measured EWs are quite large, as suggested, for example, in Mrk 766 by variability studies (Miller et al. 2006).

Both the Fe XXV and Fe XXVI lines are detected in 24 sources, 15 broad-line objects and 6 narrow-line objects. We plot in Fig. 8 the EW of one line against the other. The ratio between the two EWs is in agreement with photoionised models, but on average pointing towards large values of the ionization parameter (Bianchi & Matt 2002). This is not surprising, given the fact that these sources are those where both lines are observed and thus the gas must be ionised enough to produce a large EW for H-like iron. Column densities up to $\approx 10^{23}$ cm $^{-2}$ are required to reproduce most of the observed EWs for He-like iron, while the column densities suggested by the Fe XXVI EWs are, on average,

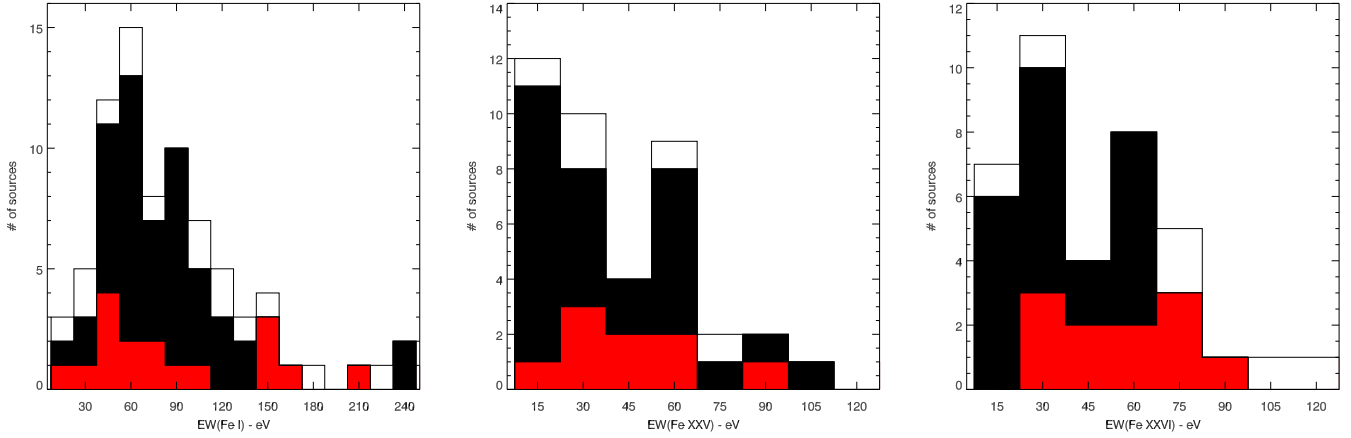


Fig. 7. Distribution of neutral iron, Fe XXV (*left*) and Fe XXVI (*right*) measured (no upper limits) EWs in CAIXA. Red histograms are narrow-line objects, black histograms are broad line objects and white histograms are for objects with no measure in CAIXA for the $H\beta$ FWHM. Very large EWs are not included in these distributions. See text for details.

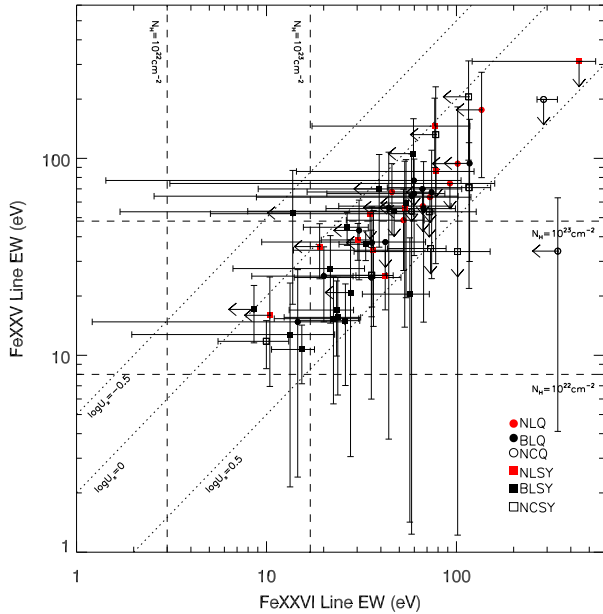


Fig. 8. Fe XXV against Fe XXVI EWs for the objects in CAIXA where at least one of the emission lines is detected (upper limits for both lines are excluded). The diagonal lines show the expected ratios between the two for three different values of the X-ray ionization parameter (see Bianchi & Matt 2002, for details). The dashed lines show the EWs expected for the two lines if they are produced in a gas photoionised by the AGN (see Bianchi & Matt 2002, for details). The different symbols refer to the classification of the objects, on the basis of their absolute magnitude and $H\beta$ FWHM: *NLSY*, narrow-line Seyfert 1; *BLSY*, broad-line Seyfert 1; *NCSY*, not-classified Seyfert 1 (no $H\beta$ FWHM measure available); *NLQ*, narrow-line quasar; *BLQ*, broad-line quasar; *ONCQ*, not-classified quasar (no $H\beta$ FWHM measure available).

larger. This result is not uncommon: see Bianchi et al. (2005) for a detailed discussion.

3.4. The soft excess

As mentioned in Sect. 3.1, the presence of a second powerlaw to model the soft X-ray emission is statistically required in 123 spectra out of 156. Although the need for this component is clearly requested in the presence of neutral absorption of the primary continuum (which is present only in 14 sources), in the majority of the cases it reflects the presence of an excess in the soft X-ray regime with respect to the intrinsic powerlaw emission. Considering also that most of the single-powerlaw fits are clustered in the low-SNR spectra, where the presence of a soft excess is more difficult to detect, we can conclude that this component is very common in CAIXA, in agreement with previous studies (see e.g. Piconcelli et al. 2004).

The modelisation of the soft excess with a simple powerlaw has no physical justification. Indeed, the values of the soft X-ray indexes found in CAIXA are often very steep (see Table 2) and cannot be naturally explained. Moreover, as already showed by several authors (Gierliński & Done 2004; Crummy et al. 2006), the soft excess in unobscured radio-quiet AGN, when fitted with a thermal model, is characterized by a temperature which is amazingly constant across orders of magnitude of luminosities and BH masses. We tested the same scenario in CAIXA, adopting another fitting model, with a thermal black body component instead of the soft powerlaw. The results are shown in Fig. 9: even on a catalogue of more sources and which spans larger orders of magnitude in luminosities and BH masses than previous studies, the temperature of the thermal emission appears completely unrelated to the fundamental properties of the AGN. Moreover, it was shown that, even for the same object, a variation of the soft X-ray luminosity of one order of magnitude does not produce any variation of this ‘universal’ temperature, in contrast with what expected from the basic physics of disk accretion and black bodies (Ponti et al. 2006).

In the last few years, two main competing models have been proposed to explain the origin of the soft excess, both ascribing the apparent universal temperature of the soft excess to atomic physics processes. In one scenario, the soft excess arises from the enhancement of reflection from the inner regions of the accretion disk due to light-bending effects, together with a strong suppression of the primary emission (e.g. Miniutti & Fabian 2004). On the other hand, the soft excess can also be mimicked

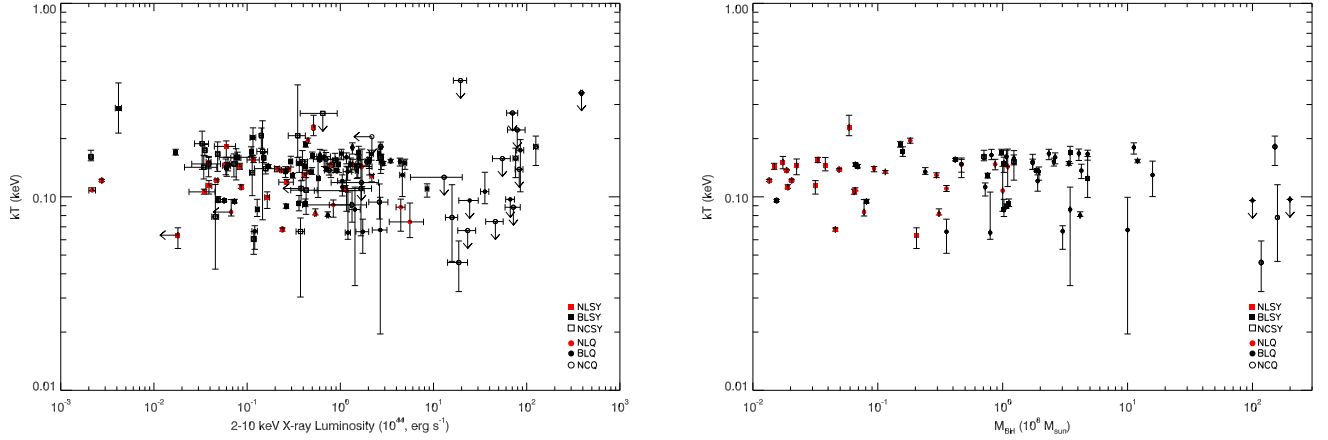


Fig. 9. Temperature of the black body model, when used to model the soft excess in the objects of CAIXA. It appears completely unrelated to the X-ray luminosity (*left*) and the BH mass (*right*). The different symbols refers to the classification of the objects, on the basis of their absolute magnitude and H β FWHM: *NLSY*, narrow-line Seyfert 1; *BLSY*, broad-line Seyfert 1; *NCSY*, not-classified Seyfert 1 (no H β FWHM measure available); *NLQ*, narrow-line quasar; *BLQ*, broad-line quasar; *NCQ*, not-classified quasar (no H β FWHM measure available).

by absorption from a relativistically outflowing warm gas (e.g. Gierliński & Done 2004). Present X-ray spectra do not allow us to discriminate between the two models.

Therefore, we decided to investigate the nature of the soft excess adopting a model-independent parameter. We chose the ratio between the soft X-ray luminosity (rest-frame 0.5-2 keV) and the hard X-ray luminosity (rest-frame 2-10 keV) as an indicator of the importance of the soft excess with respect to the overall X-ray spectrum of the source. As shown in Sect. 3.1, the luminosities extracted with the automatic fitting procedure in CAIXA are not significantly affected by the adopted model and so their ratio is a good parameter to characterize at least the basic property of the soft excess of the sources, i.e. its strength with respect to the primary X-ray emission.

In Fig. 10 we show the distribution of the X-ray luminosity ratio in CAIXA. As already mentioned in Sect. 2.1, we stress here that this value is calculated only for sources with no neutral absorption, for which the intrinsic soft X-ray luminosity is not measurable. On the other hand, as already commented in previous sections, a warm absorber component is not modeled in CAIXA, so some of the calculated X-ray luminosity ratios may be lower than the intrinsic ones: in some objects, low ratios may not be an intrinsic property, but instead they may indicate the presence of a strong warm absorption component. Although the majority of the sources cluster around a value of 1, narrow- and broad-line objects are significantly different populations with respect to the X-ray luminosity ratio, the NHP being 6×10^{-9} for a K-S test. Indeed, it is very interesting to note that no narrow-line object is found with a value of the luminosity ratio lower than unity. In other words, narrow-line objects tend to have a strong soft X-ray luminosity with respect to their hard X-ray emission. We defer the reader to a forthcoming companion CAIXA paper for a detailed discussion on the correlations between the X-ray luminosity ratio and other multiwavelength properties of AGN, and its possible consequences on the origin of the soft excess.

4. Conclusions

We presented CAIXA, a Catalogue of AGN In the XMM-Newton Archive. It consists of all the radio-quiet X-ray unob-

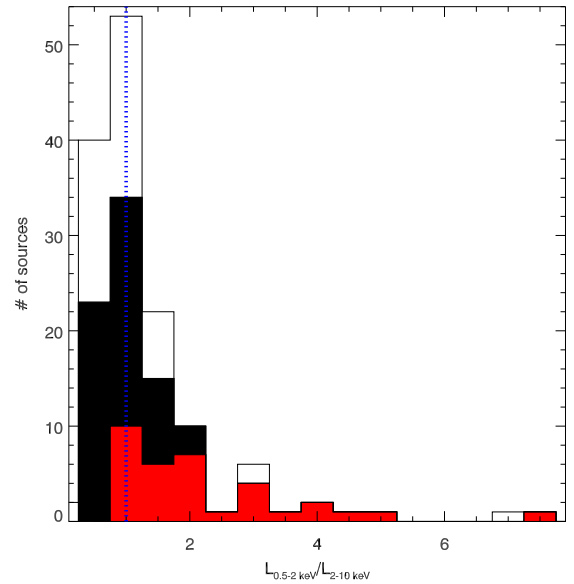


Fig. 10. Hard X-ray powerlaw index distribution in CAIXA. Red histograms are narrow-line objects, black histograms are broad line objects and white histograms are for objects with no measure in CAIXA for the H β FWHM. See text for details.

scured ($N_H < 2 \times 10^{22} \text{ cm}^{-2}$) AGN observed by XMM-Newton in targeted observations, whose data are public as of March 2007. We performed a complete and homogeneous spectral analysis of the X-ray data, whose main results can be summarised as follows:

- *The X-ray spectral index.* The average 2-10 keV spectral index for the whole catalogue is 1.73 ± 0.04 ($\sigma = 0.49 \pm 0.02$) and appears significantly steeper in narrow-line objects and quasars with respect to broad-line objects and Seyfert galaxies. The average value for CAIXA and the differences between source populations do not significantly change if a

luminosity-dependent Compton reflection component is included in the model.

- *The iron lines.* The distribution of the neutral iron $K\alpha$ EW is clustered around 50-120 eV, as expected if arising as reprocessing from cold Compton-thick matter, as the putative torus. Fe XXV and Fe XXVI are commonly detected, but we do not find different rates of detection of these lines between narrow- and broad-line objects. The ratio between the EWs of the two lines is, on average, broadly in agreement with photoionisation models.
- *The soft excess.* This component is very common in CAIXA, in agreement with previous studies. We chose the ratio between the soft X-ray luminosity (rest-frame 0.5-2 keV) and the hard X-ray luminosity (rest-frame 2-10 keV) as a model-independent indicator of the strength of the soft excess. With respect to this parameter, narrow- and broad-line objects are significantly different populations. Indeed, it is very interesting to note that no narrow-line object is found with a value of the luminosity ratio lower than unity.

On the basis of these results, we will investigate the correlations between the X-ray and the multiwavelength properties of the sources in CAIXA in a companion paper.

Acknowledgements. SB and GM acknowledge financial support from ASI (grant I/088/06/0), GP from ANR (ANR-06-JCJC-0047). We acknowledge support from the Faculty of the European Space Astronomy Centre (ESAC). We thank C. Gordon for all his efforts in solving problems related to XSPEC and the anonymous referee for his/her suggestions. Based on observations obtained with XMM-Newton, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA.

Appendix A: Censored mean

Given the large errors and the presence of upper limits for many of the parameters derived in CAIXA, most average values cited in this paper are ‘censored means’, if not otherwise stated. The adopted method is very similar to the ‘censored fits’ presented in Bianchi et al. (2007), based on the linear fits performed by Guainazzi et al. (2006b), and can be summarized as follows. A large number of mean values was performed on a set of Monte-Carlo simulated data derived from the experimental points according to the following rules: a) each detection was substituted by a random Gaussian distribution, whose mean is the best-fit measurement and whose standard deviation is its statistical uncertainty; b) each upper limit U was substituted by a random uniform distribution in the interval $[0, U]$. The average (with relative uncertainty) of the mean values of each single data set is the ‘censored mean’ reported in the text, along with the average dispersion.

References

- Anders, E. & Ebihara, M. 1982, *Geochim. Cosmochim. Acta*, 46, 2363
- Arnaud, K. A. 1996, in ASP Conf. Ser. 101: *Astronomical Data Analysis Software and Systems V*, 17
- Bian, W. & Zhao, Y. 2003a, *MNRAS*, 343, 164
- Bian, W. & Zhao, Y. 2003b, *MNRAS*, 343, 164
- Bian, W. & Zhao, Y. 2004, *MNRAS*, 352, 823
- Bianchi, S., Guainazzi, M., Matt, G., & Fonseca Bonilla, N. 2007, *A&A*, 467, L19
- Bianchi, S. & Matt, G. 2002, *A&A*, 387, 76
- Bianchi, S., Matt, G., Nicastro, F., Porquet, D., & Dubau, J. 2005, *MNRAS*, 357, 599
- Boller, T., Brandt, W. N., & Fink, H. 1996, *A&A*, 305, 53
- Boroson, T. A. & Green, R. F. 1992, *ApJS*, 80, 109
- Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, *AJ*, 115, 1693
- Corbin, M. R. 1991, *ApJ*, 371, L51
- Corbin, M. R. & Boroson, T. A. 1996, *ApJS*, 107, 69
- Crenshaw, D. M., Kraemer, S. B., & Gabel, J. R. 2003, *AJ*, 126, 1690
- Crummy, J., Fabian, A. C., Gallo, L., & Ross, R. R. 2006, *MNRAS*, 365, 1067
- Czerny, B., Rózańska, A., & Kuraszkiewicz, J. 2004, *A&A*, 428, 39
- Dadina, M. 2008, *ArXiv e-prints*, 801
- Dickey, J. M. & Lockman, F. J. 1990, *ARA&A*, 28, 215
- Gabriel, C., Denby, M., Fyfe, D. J., et al. 2004, in ASP Conf. Ser. 314: *Astronomical Data Analysis Software and Systems (ADASS) XIII*, 759
- Gallo, L. C., Boller, T., Brandt, W. N., Fabian, A. C., & Grupe, D. 2004, *MNRAS*, 352, 744
- Ganguly, R., Bond, N. A., Charlton, J. C., et al. 2001, *ApJ*, 549, 133
- Ganguly, R., Sembach, K. R., Tripp, T. M., Savage, B. D., & Wakker, B. P. 2006, *ApJ*, 645, 868
- Gierliński, M. & Done, C. 2004, *MNRAS*, 349, L7
- Gregory, P. C., Scott, W. K., Douglas, K., & Condon, J. J. 1996, *ApJS*, 103, 427
- Grupe, D., Beuermann, K., Mannheim, K., & Thomas, H.-C. 1999, *A&A*, 350, 805
- Grupe, D., Wills, B. J., Leighly, K. M., & Meusinger, H. 2004, *AJ*, 127, 156
- Guainazzi, M., Bianchi, S., & Dovčiak, M. 2006a, *Astronomische Nachrichten*, 327, 1032
- Guainazzi, M., Siemiginowska, A., Stanghellini, C., et al. 2006b, *A&A*, 446, 87
- Kellermann, K. I., Sramek, R., Schmidt, M., Shaffer, D. B., & Green, R. 1989, *AJ*, 98, 1195
- Kuhn, O., Elvis, M., Bechtold, J., & Elston, R. 2001, *ApJS*, 136, 225
- Laor, A. 1998, *ApJ*, 505, L83
- Laor, A. 2000, *ApJ*, 543, L111
- Leighly, K. M. 1999, *ApJS*, 125, 317
- Longinotti, A. L., de La Calle, I., Bianchi, S., Guainazzi, M., & Dovčiak, M. 2007, *ArXiv e-prints*, 709
- Marconi, A., Risaliti, G., Gilli, R., et al. 2004, *MNRAS*, 351, 169
- Marziani, P., Sulentic, J. W., Zamanov, R., et al. 2003, *ApJS*, 145, 199
- Matt, G., Perola, G. C., & Piro, L. 1991, *A&A*, 247, 25
- Mauch, T., Murphy, T., Buttery, H. J., et al. 2003, *MNRAS*, 342, 1117
- McIntosh, D. H., Rieke, M. J., Rix, H.-W., Foltz, C. B., & Weymann, R. J. 1999, *ApJ*, 514, 40
- McLure, R. J. & Dunlop, J. S. 2001, *MNRAS*, 327, 199
- McLure, R. J. & Dunlop, J. S. 2004, *MNRAS*, 352, 1390
- McMahon, R. G., Priddey, R. S., Omont, A., Snellen, I., & Withington, S. 1999, *MNRAS*, 309, L1
- Merloni, A., Heinz, S., & di Matteo, T. 2003, *MNRAS*, 345, 1057
- Miller, L., Turner, T. J., Reeves, J. N., et al. 2006, *A&A*, 453, L13
- Miniutti, G. & Fabian, A. C. 2004, *MNRAS*, 349, 1435
- Molendi, S., Bianchi, S., & Matt, G. 2003, *MNRAS*, 343, L1
- Netzer, H. & Trakhtenbrot, B. 2007, *ApJ*, 654, 754
- O’Neill, P. M., Nandra, K., Papadakis, I. E., & Turner, T. J. 2005, *MNRAS*, 358, 1405
- Page, K. L., O’Brien, P. T., Reeves, J. N., & Turner, M. J. L. 2004, *MNRAS*, 347, 316
- Panessa, F., Barcons, X., Bassani, L., et al. 2007, *astro-ph/0701546*
- Peterson, B. M., Ferrarese, L., Gilbert, K. M., et al. 2004, *ApJ*, 613, 682
- Piconcelli, E., Jimenez-Bailón, E., Guainazzi, M., et al. 2004, *MNRAS*, 351, 161
- Piconcelli, E., Jimenez-Bailón, E., Guainazzi, M., et al. 2005, *A&A*, 432, 15
- Ponti, G., Miniutti, G., Cappi, M., et al. 2006, *MNRAS*, 368, 903
- Read, A. M. & Ponman, T. J. 2003, *A&A*, 409, 395
- Reeves, J. N., Nandra, K., George, I. M., et al. 2004, *ApJ*, 602, 648
- Reynolds, C. S., Ward, M. J., Fabian, A. C., & Celotti, A. 1997, *MNRAS*, 291, 403
- Sadler, E. M., Jackson, C. A., Cannon, R. D., et al. 2002, *MNRAS*, 329, 227
- Sadler, E. M., Ricci, R., Ekers, R. D., et al. 2006, *MNRAS*, 371, 898
- Shemmer, O., Brandt, W. N., Netzer, H., Maiolino, R., & Kaspi, S. 2006, *ApJ*, 646, L29
- Smith, M. G. & Wright, A. E. 1980, *MNRAS*, 191, 871
- Stoeck, J. T., Morris, S. L., Weymann, R. J., & Foltz, C. B. 1992, *ApJ*, 396, 487
- Strüder, L., Briel, U., Dennerl, K., et al. 2001, *A&A*, 365, L18
- Terashima, Y. & Wilson, A. S. 2003, *ApJ*, 583, 145
- Veron-Cetty, M. P. & Veron, P. 2006, *VizieR Online Data Catalog*, 7248, 0
- Veron-Cetty, M. P., Woltjer, L., & Roy, A. L. 1991, *A&A*, 246, L73
- Vestergaard, M. 2002, *ApJ*, 571, 733
- Vestergaard, M. & Peterson, B. M. 2006, *ApJ*, 641, 689
- Wandel, A., Peterson, B. M., & Malkan, M. A. 1999, *ApJ*, 526, 579
- Wang, J.-M., Watarai, K.-Y., & Mineshige, S. 2004, *ApJ*, 607, L107
- Wang, T., Brinkmann, W., & Bergeron, J. 1996, *A&A*, 309, 81
- Wang, T. & Lu, Y. 2001, *A&A*, 377, 52
- Winkler, H. 1992, *MNRAS*, 257, 677
- Woo, J. & Urry, C. M. 2002, *ApJ*, 579, 530
- Yuan, W., Brinkmann, W., Siebert, J., & Voges, W. 1998, *A&A*, 330, 108
- Zhang, E.-P. & Wang, J.-M. 2006, *ApJ*, 653, 137
- Zhou, X.-L. & Wang, J.-M. 2005, *ApJ*, 618, L83

Table 1. Main properties of the observations of the sources included in CAIXA.

Source	ObsID	T (ks)	F(0.5-2 keV)	F(2-10 keV)	L(0.5-2 keV)	L(2-10 keV)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
..

Full table will be published in electronic form at the CDS.

(1) Source name (Veron-Cetty & Veron 2006); (2) XMM-Newton OBSID (3) Clean exposure time; (4) 0.5-2 keV flux (10^{-12} erg cm $^{-2}$ s $^{-1}$); (5) 2-10 keV flux (10^{-12} erg cm $^{-2}$ s $^{-1}$); (6) Unabsorbed 0.5-2 keV luminosity (10^{44} erg s $^{-1}$); (7): Unabsorbed 2-10 keV luminosity (10^{44} erg s $^{-1}$).

Table 2. Main X-ray parameters derived from the automatic fits of the observations included in CAIXA.

Source	Obsid	Γ_s	Γ_h	EW_{6.4}	EW_{6.7}	EW_{6.96}
(1)	(2)	(3)	(4)	(5)	(6)	(7)
..

Full table will be published in electronic form at the CDS.

(1) Source name (Veron-Cetty & Veron 2006); (2) XMM-Newton OBSID; (3) Soft X-ray powerlaw index; (4) Hard X-ray powerlaw index; (5) 6.4 keV emission line EW (eV); (6) 6.7 keV emission line EW (eV); (7) 6.96 keV emission line EW (eV).

Table 3. Multiwavelength properties of the sources included in CAIXA. RQ means the source is considered radio-quiet in the reference, but no radio flux is reported.

NAME	RA	DEC	z	M _{abs}	F _{6cm}	F _{20cm}	H _β	M _{BH}	m _v	B-V
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
..

Full table will be published in electronic form at the CDS.

(1) Source name (Veron-Cetty & Veron 2006); (2)-(3): Coordinates in J2000 (NED); (4) Redshift (Veron-Cetty & Veron 2006); (5) Absolute magnitude (Veron-Cetty & Veron 2006); (6)-(7): Radio fluxes at 6 and 20 cm (Jy). References: 1. Veron-Cetty & Veron (2006), 2. Kellermann et al. (1989), 3. Gregory et al. (1996), 4. Condon et al. (1998), 4u. Sensitivity limit in Condon et al. (1998), 5. Kuhn et al. (2001), 6. Veron-Cetty et al. (1991), 7. Gallo et al. (2004), 8. Smith & Wright (1980), 9. Page et al. (2004), 10. Ganguly et al. (2006), 11. Mauch et al. (2003), 12. Yuan et al. (1998), 13. Wang et al. (2004) 14. Sadler et al. (2002), 15. , 16. McMahon et al. (1999); (8) FWHM (km s⁻¹). References: a. Boroson & Green (1992), b. Wang & Lu (2001), c. Vestergaard (2002), d. Wandel et al. (1999), e. Boller et al. (1996), f. Grupe et al. (1999) g. Grupe et al. (2004), h. Bian & Zhao (2003a), i. Peterson et al. (2004), j. Wang et al. (1996), k. Leighly (1999), l. Marziani et al. (2003), m. McIntosh et al. (1999), n. Winkler (1992), o. Ganguly et al. (2001), p. Corbin (1991), q. Corbin & Boroson (1996), r. Shemmer et al. (2006), s. Crenshaw et al. (2003), t. McLure & Dunlop (2001), u. Laor (2000), v. Reynolds et al. (1997), w. Zhou & Wang (2005); (9) Log of BH mass in units of solar masses. References: A. Woo & Urry (2002), B. Vestergaard & Peterson (2006), C. Vestergaard (2002), D. Zhang & Wang (2006), E. Wang & Lu (2001), F. Shemmer et al. (2006), G. Czerny et al. (2004), H. Laor (1998), I. Bian & Zhao (2003b), J. O'Neill et al. (2005), K. Merloni et al. (2003), L. Zhou & Wang (2005) - Notes *r*:reverberation mapping, *o*: H_β FWHM, *m*: MgII FWHM, *v*: stellar velocity dispersion; *n*: narrow line region size (10) *m_v* (Veron-Cetty & Veron 2006); (11) B-V (mag) (Veron-Cetty & Veron 2006).